Preventive Veterinary Medicine 89 (2009) 43–50

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

journal homepage: www.elsevier.com/locate/prevetmed

Use of epidemiologic information in targeted surveillance for population inference

Scott J. Wells [a,*], Eric D. Ebel [b], Michael S. Williams [b], Aaron E. Scott [c], Bruce A. Wagner [d], Katherine L. Marshall [d]

[a] Department of Veterinary Population Medicine, University of Minnesota, St. Paul, MN 55108, USA
[b] Risk Analysis Division, Food Safety Inspection Service, USDA, Ft. Collins, CO 80526, USA
[c] National Surveillance Unit, Centers for Epidemiology and Animal Health, USDA-APHIS-Veterinary Services, Ft. Collins, CO 80526, USA
[d] National Animal Health Monitoring System, Centers for Epidemiology and Animal Health, USDA-APHIS-Veterinary Services, Ft. Collins, CO 80526, USA

ARTICLE INFO

Article history:
Received 14 January 2008
Received in revised form 14 January 2009
Accepted 20 January 2009

Keywords:
Targeted surveillance
Prevalence estimation
Risk ratio

ABSTRACT

Epidemiologic information, including animal characteristics (e.g., observable risk factors or clinical signs) predisposing to animal disease, is frequently used for design of targeted surveillance systems, but this information is infrequently used for population inference. In this study, we report the evaluation of use of epidemiologic information for population inference in targeted surveillance in three animal disease scenarios. We adapted sampling theory using Monte Carlo methods to determine target population sample size to detect disease with 95% confidence, using information from the epidemiologic parameters risk ratio and fraction of the population with the characteristic. These parameters and their uncertainties were derived from a reference population. The next step was to use a second (sampled) population to evaluate effects of sampling the targeted population. The focus of the study was on estimation of prevalence. Our results showed that if one is less certain of the epidemiologic parameters, a rational decision is to model the input parameter distributions reflecting this uncertainty, thereby increasing the sample size above the minimum needed for the detection of the disease with a known confidence. Targeted surveillance is appropriate for prevalence estimation when one has representative and justifiable estimates of key epidemiologic parameters.

1. Introduction

Targeted sampling is emerging as an important concept in animal health surveillance. Classically, one type of targeted surveillance has been visual observation of animals within herds with specific clinical signs of disease. This form of targeted surveillance has been used historically in veterinary medicine for case finding and has proven critical for detection and eradication of important diseases including foot and mouth disease, contagious pleuropneumonia, and others. The past 5 years, however, have seen a substantial effort by many countries to incorporate targeted sampling concepts into eradication programs for endemic diseases and for demonstrating freedom from disease (e.g., Stärk et al., 2006). The US Bovine Spongiform Encephalopathy (BSE) Surveillance Program has used targeted surveillance (BSurvE) to enhance the likelihood of detection of BSE in cattle with clinical signs of disease or other abnormal presentations (USDA-APHIS-VS, 2006). In this surveillance system, a complex model involving age and categories for exiting the population was used to derive point values for each

* Corresponding author at: 136 ABLMS, 1354 Eckles Avenue, University of Minnesota, St. Paul, MN 55108, USA. Tel.: +1 612 625 8166; fax: +1 612 624 4906.
E-mail address: wells023@umn.edu (S.J. Wells).

0167-5877/$ – see front matter © 2009 Published by Elsevier Ltd.
doi:10.1016/j.prevetmed.2009.01.007

sampling stratum (Prattley et al., 2007). Recently, the Bovine Tuberculosis Surveillance Program in Minnesota has focused sampling on cattle herds with suspected risk factors for infected herd status, including history of cattle introductions from the southwest US or exposure to roping steers and proximity to known infected herds (Minnesota Board of Animal Health, 2006). Similarly, the US Scrapie Surveillance Program focuses sampling on black-faced adult sheep due to the increased risk of infection among this targeted group (USDA-APHIS-VS, 2004). Much of the targeted surveillance work to date has been geared towards specific disease applications. The result of these efforts is a collection of analytic approaches that tend to be application specific and usually not used for population inference.

The general concept of targeted sampling is to choose samples from the population with probabilities based on a characteristic, defined as an easily observed clinical sign or other diagnostic factor. In these applications, the presence of the characteristic indicates an increased probability for the presence of disease. Thus, targeted sampling is a form of unequal probability sampling. Targeted sampling is attractive for surveillance because unequal probability sampling often results in a substantial reduction in the required sample size through focusing on a subpopulation with higher disease prevalence. Targeted surveillance therefore can improve the economic efficiency of surveillance compared to simple random sampling.

Results from targeted sampling appear to have been derived without considering the connections between targeted sampling and standard finite population sampling theory. For example, results from targeted surveillance have been used predominantly for disease detection and not for prevalence estimation of animal disease. One factor that has led epidemiologists to this departure is the cost and difficulty of constructing a sampling frame. Another factor is that sample collection tends to be opportunistic, which leads to inference based on an assumed model, rather than the sample design (Särndal, 1978). While there is nothing inherently wrong with opportunistic sampling and model-based inference, there is little acknowledgement of the connections between traditional sampling theory and targeted sampling. Therefore, there is rarely an appreciation of the potential application of standard survey sampling theory to targeted sampling based on an epidemiologic model.

A mathematical theory for targeted sampling has been previously described (Williams et al., 2009). The fundamental principle of targeted surveillance is to focus sampling on one or more subpopulations with higher prevalence. This sampling focus allows use of a smaller sample size for surveillance. A characteristic (e.g., an observable risk factor or clinical sign) divides the population. To estimate the sample size required for detection using a risk-based system, a relative value, or points, is associated with each subpopulation (Cannon, 2002; Prattley et al., 2007). An animal randomly selected from the general population has a point value of 1; the elevated point value for the subpopulation of interest reflects the increased likelihood that samples from this subpopulation will be from infected/diseased individuals. The mathematics for targeted sampling is based on the risk ratio of the characteristic, the fraction of the population with the characteristic, and a population or design prevalence. Knowledge of these three parameters is sufficient to design, implement, and analyze data from a targeted surveillance application. Nevertheless, practical concerns about targeted surveillance stem from uncertainties surrounding risk ratio estimation, as well as estimates of the fraction of the population with the characteristics. These uncertainties include those defined by statistical methods as well as the uncertainties surrounding the validity of measurements from the reference population (in space and time) as it relates to the population to be sampled.

The objective of this study was to evaluate the performance of targeted sampling in specific examples of animal disease surveillance. This work is built on theory previously developed (Williams et al., 2009) by applying those methods to real world situations. One outcome of this study was a transparent description of the necessary steps for designing and analyzing data from a targeted surveillance application. This outcome should support future surveillance efforts. Better understanding of targeted surveillance will improve national capabilities related to monitoring and surveillance of economically important disease conditions in livestock.

2. Methods

2.1. Study design

A targeted subpopulation is defined as a segment of the population with higher disease prevalence than the general population and exhibiting a specific characteristic. In this paper we demonstrate the design and analysis of targeted sampling for three animal disease scenarios. These scenarios were chosen because each differs with respect to the validity of risk ratio and population fraction when applied to the sampled population (i.e., these parameters are progressively less representative of the populations that are sampled). In each case, data on the prevalence of disease, risk ratio of a characteristic predisposing to disease, and the prevalence of this characteristic were derived from a *reference* population and used in conjunction with Monte Carlo methods to determine the necessary sample size for a targeted surveillance program. Both SRS and targeted surveillance samples were then drawn from a second population, which will be referred to as the *sampled population*, to estimate both the prevalence of disease and the probability of disease detection. Test sensitivity and specificity were assumed to be 100% to simplify these examples, resulting in apparent prevalence values equivalent to true prevalence values.

Specifically, the prevalence, risk ratio of the characteristic, and fraction of the population with the characteristic were obtained for three animal disease scenarios (scrapie among sheep, Johne's disease among dairy cattle, and *Neospora caninum* among dairy cattle) from National Animal Health Monitoring Systems (NAHMS) studies (USDA-APHIS-Veterinary Services). Estimates of the

most-likely (in this case the mean) value and 95% confidence intervals for risk ratio were fit to pert distributions using RiskView (@Risk, Risk Analysis Add-in for Microsoft Excel, Palisade Corp., Ithaca, NY, 2004). Similarly, estimates for fraction of the target population were fit to beta distributions using BetaBuster (Betabuster 1.0, http://www.epi.ucdavis.edu/diagnostictests/betabuster.html).

Design prevalence levels were assumed to be the apparent (observed) prevalence of the condition determined from reference populations. The size of a simple random sample needed to detect infection in a large population was

$$n = \frac{\ln(\alpha)}{\ln(1-P)}. \quad (1.1)$$

This prevalence, P, is the *design prevalence*, which is a user-defined threshold that plays a key role in determining the sample size required to detect the disease. The sampling error to determine the confidence in detecting one or more positive animals in a sample from an infected population (α) was assigned as 0.05.

Targeted sampling calculations were previously derived (Williams et al., 2009). Briefly, the number of points per targeted sample (γ) was assumed to equal P_T/P, where P_T was prevalence in the target subpopulation. Algebra was used to derive $\tilde{\gamma} = R\tilde{R}/(\tilde{f}_T R\tilde{R} + (1-\tilde{f}_T))$ where RR was the risk ratio, f_T was the fraction of the population with the characteristic, and the $\sim$ symbol indicated that all quantities were random variables derived from a specified distribution. To determine the number of targeted samples (n_T), 10,000 instances of

$$\tilde{P}_T = \tilde{\gamma}P = \frac{PR\tilde{R}}{\tilde{f}_T R\tilde{R} + (1-\tilde{f}_T)} \quad (1.2)$$

were generated from Monte Carlo simulation. Using the simulated values for P_T, the value for n_T that solves

$$\frac{\sum_{i=1}^{10,000}(1-(1-\tilde{P}_{T_i})^{n_T})}{10,000} \approx 0.95 \quad (1.3)$$

was determined. The solver utility in Excel was used.

To compare the performance of targeted sampling with simple random sampling, Monte Carlo simulation was performed using commercial software (@Risk, Risk Analysis Add-in for Microsoft Excel, Palisade Corp., Ithaca, NY, 2004) by sampling with replacement for 20,000 iterations. The outcomes of interest were the mean estimated prevalence and the standard error (SE) of the estimated prevalence. Using simple random sampling, the estimated prevalence estimator was $\hat{P} = s/n$ and the SE of prevalence was, $SE = \sqrt{\hat{P}(1-\hat{P})/n}$ where s was the number of positive animals in the sampled population. For targeted sampling, the estimated prevalence estimator was $\hat{P} = s_T/(\gamma n_T)$ and the SE of prevalence was $SE = \sqrt{\hat{P}(1-\hat{P})/(\gamma n_T)}$, where s_T was the number of positive animals in the target sample.

The bias ratio was calculated as the difference between the estimated prevalences using unbiased (simple random sampling) and biased (targeted sampling) estimators divided by the standard error of the biased prevalence estimate. The motivation for using the bias ratio is that this statistic can be used to assess the acceptable level of bias in an estimator. Särndal et al. (1992, pp. 164–166) demonstrate that a bias ratio less than 0.50 is often times acceptable because the achieved coverage probability of the confidence interval is not greatly affected by the bias (i.e., the proportion of samples with 95% confidence intervals that overlap the true prevalence is still reasonably close to 95%). For example, Särndal et al. (1992) show that when the bias ratio is 0.5, the achieved coverage rate of a 95% confidence interval is still 92%. For bias ratio values greater than 0.5, the coverage probabilities correspondingly decrease and the inferences drawn from biased estimators become progressively less valid.

2.2. Disease scenarios (Table 1)

2.2.1. Scrapie in US sheep

This scenario compares the performance of simple random and targeted sampling strategies in the case where a highly representative reference population has been used to estimate RR and f_T prior to the design phase of surveillance (i.e., the reference population is representative of the population to be sampled). A national slaughter study was performed in 2002–2003 to estimate the prevalence of scrapie histopathological lesions in the US adult sheep population using immunohistochemistry (IHC) of the obex, tonsil, and lymph node (USDA-APHIS-VS, 2004), the standard diagnostic test performed for this disease in US sheep.

A strong association between black-faced sheep and occurrence of scrapie has long been assumed (USDA-APHIS-VS, 2003). For our study, we used a target population consisting of black-faced sheep. To create a situation of highly representative input data, the national study dataset was randomly split into two equal sized datasets irrespective of face color, thereby providing reference and sampled populations derived from the same population. From the reference dataset, estimates of RR, f_T, and P were obtained (USDA-APHIS-VS, 2004). Uncertainty distributions for RR and f_T implied by the estimates from this dataset were used to derive the targeted sample size. The prevalence, P, from the reference dataset was then chosen as the design prevalence for the calculation of the required sample size necessary for both simple random sampling (Eq. (1.1)) and targeted sampling (Eq. (1.2)).

In the sampled dataset, the exact values were determined for the prevalence of scrapie among black-faced sheep, the fraction of the population that were black-faced, and the prevalence of scrapie among all sheep in the population. These population parameters were used in a simulation study to compare the performance of the targeted sampling prevalence estimator with the simple random sampling estimator. Using Monte Carlo methods, bootstrap sampling was employed to calculate the estimated value and standard error of prevalence using both sampling approaches. For example, bootstrap simple random sampling was accomplished by simulating 20,000 iterations of $\tilde{s}/n$ where $\tilde{s}$ is a binomial variable ($s \sim Binomial(n, P)$), n is the simple random sample size and P is the prevalence of scrapie in the sampled

Table 1
Data sources and parameter estimates for specific disease situations.

Disease	Characteristic	Risk ratio (*RR*) (95% CI)	Fraction of population with characteristic (f_T) (95% CI)	Prevalence in target population (P_T)	Overall prevalence (P)
Scrapie in US sheep					
Reference population	Black face	20.02% (4.49–89.33)	23.06% (22.01–24.12%)	–	0.23%
Sampled population	Black face	26.75%	21.90%	1.09%	0.27%
Johne's disease in US dairy cattle					
Reference population	Watery diarrhea	4.66% (2.60–8.37)	1.86% (1.71–2.01%)	–	0.47%
Sampled population	Watery diarrhea	2.75%	0.62%	5.08%	1.87%
Neospora caninum in US dairy cattle					
Reference population	Abortion	2.80% (1.10-6.90)	3.50% (3.30-3.70%)	–	1.00%
Sampled population	Abortion	4.79%	3.25%	2.47%	0.58%

Sources of information:
Scrapie in US sheep – Reference population: SOSS 2002–2003 Part A. Sample size = 6205. Diagnosis by IHC of obex, tonsil, or lymph node; and Sampled population: SOSS 2002–2003 Part B. Sample size = 6211. Diagnosis by IHC of obex, tonsil, or lymph node.
Johne's disease in US dairy cattle – Reference population: NAHMS 1996. Sample size = 31,345. Diagnosis by serum ELISA with cutpoint OD $\geq$0.35 above mean negative control OD; and Sampled population: NAHMS 2002. Sample size = 11,151. Diagnosis by serum ELISA with cutpoint OD $\geq$0.35 above mean negative control OD.
Neospora caninum in US dairy cattle – Reference population: Hernandez et al. (2002) for RR estimate. Sample size = 460. USDA-APHIS-VS (1996), Part I: Reference of 1996 Dairy Management Practices for f_T; and Sampled population: NAHMS 1996. Sample size = 7224. Diagnosis in 2nd lactation dairy cattle by serum ELISA with cutpoint OD $\geq 1.2 \times$ high positive control.

population. The estimated value and standard deviation of $\tilde{s}/n$ across the 20,000 iterations approximates the estimated prevalence from simple random sample and its standard error. For targeted sampling, a similar procedure is followed but in this case 20,000 iterations of $\tilde{s}_T/\tilde{\gamma}n_T$ are simulated where $s_T \sim Binomial(n_T, P_T)$, n_T is the targeted sample size, P_T is the prevalence of scrapie among black-faced sheep in the sampled population and $\tilde{\gamma}$ is a random variable whose distribution is predicted by $R\tilde{R}/(\tilde{f}_T R\tilde{R} + (1 - \tilde{f}_T))$.

The estimated prevalence and its standard error are calculated as described for simple random sampling.

2.2.2. Johne's disease (*Mycobacterium avium subsp. paratuberculosis*) in US dairy cattle

This scenario compared the performance of targeted sampling and simple random sampling in the case where estimates of *RR* and f_T used to design surveillance are generated from a reference population that may not be highly representative of the current population. The reference dataset was collected in 1996 where a national study of 967 nonvaccinated dairy herds resulted in the testing of 31,745 cows for Johne's disease (USDA-APHIS-VS, 1997). Testing was performed using a serum ELISA (IDEXX, Portland, ME) at the National Veterinary Services Laboratory (NVSL), Ames, IA. Weighted results from this study, using the commercial test kit label cutpoint (0.10 above the mean negative control OD), showed an apparent cow prevalence of 2.5%, with variation observed by herd size, region, and fecal consistency score. To artificially create a sampled population with a lower disease prevalence that is more typical of situations of interest to national animal health officials, we increased the cutpoint of the test to 0.35 above the mean negative control OD, thereby reducing the apparent prevalence of infected cattle in the sample to approximately 0.5%. The characteristic was watery diarrhea as noted by the veterinarian while collecting blood samples from the cattle on the farm. These reference data were used to estimate *RR*, P and f_T, for use in the design phase of this study. The uncertainties about *RR* and f_T implied by the estimates from this dataset were used in the derivation of the targeted sample size. The observed value for P in the reference dataset was assumed to apply for the purposes of designing a simple random sampling strategy as well as a targeted sampling strategy.

To compare targeted sampling with simple random sampling, results from the NAHMS Dairy 2002 Study (USDA-APHIS-VS, 2005) were used were used to construct the sampled population. In this study, NAHMS also estimated the prevalence of antibody to *M. paratuberculosis* in dairy cows with the same diagnostic test (though marketed at this time by a different company in the US) performed at the same laboratory, though with a different study design. For our study, we used the same high cutpoint for the serum ELISA to classify cows by *M. paratuberculosis* infection status.

2.2.3. *N. caninum* in US dairy cattle

This scenario compared the performance of targeted sampling and simple random sampling in the case where estimates of *RR* and f_T used to design surveillance are generated from published, but geographically limited research that is not necessarily representative of the larger population (i.e., the reference population is not representative of the sampled population). The characteristic used to create a target subpopulation was prior history of abortion. For the design phase of our study, an estimate of *RR* for antibody to *N. caninum* due to abortion was obtained from a published study based on a single dairy herd in Florida (Hernandez et al., 2002), to simulate a situation with minimal available scientific evidence from which to design targeted surveillance. To create the sampled population, only second lactation cows were included in the dataset, as these animals were found to have increased risk of abortion related to *N. caninum* antibody in the published report. An estimate of f_T was obtained from the NAHMS Dairy 1996 Study (USDA-

APHIS-VS, 1996) as collected from producers in a retrospective herd-level questionnaire. The design prevalence was set at 1%.

To compare targeted sampling with simple random sampling, we used data from second lactation cows from the NAHMS Dairy 1996 Study, in which serum from over 30,000 cows in nearly 1000 herds was tested by NVSL in Ames, IA using an NVSL-produced ELISA to detect antibody to *N. caninum*. For this study, the cutpoint for the serum ELISA was set at an OD greater than or equal to 120% of the high positive control OD value, again to ensure low prevalence within the sample. Cattle selected for sampling had a history of aborting in the preceding 12 months at 3 or more months of gestation.

3. Results

3.1. Scrapie

The simple random sample size to detect at least one positive sheep with 95% confidence at 0.2% prevalence was 1496. The targeted sampling sample size based on most-likely values for *RR* and f_T (i.e., 20.02 and 23%) was 377, based on 3.72 points per targeted sample (Table 2). The derived targeted sample size was inflated to 405 after accounting for uncertainty in *RR* and f_T, with the mean points per targeted sample of 3.79 with 95% confidence interval from 2.49 to 4.26. These sample sizes were intended to detect one or more positive samples if the prevalence was 0.2% or greater. In the sampled population, 1.09% of the target population was scrapie-positive compared to 0.04% of the nontarget population, for an overall prevalence of 0.27%.

The required number of targeted samples was just 27% of the number of simple random samples (Table 3). Despite fewer samples, targeted sampling detected one or more positive samples slightly more often than simple random sampling (i.e., 98.8% vs. 98.3%, Table 3). The distributions of estimated prevalence for each sampling method (Fig. 1) were very similar. The standard error of estimated prevalence for simple random sampling was slightly less than that for targeted sampling and the bias ratio was 0.15.

3.2. Johne's disease

The sample size for simple random sampling in this situation was 598 (Table 2). The targeted sample size based on most-likely (i.e., fixed) values for *RR* and f_T (i.e., 4.66 and 1.86%) was 125. The derived targeted sample size was inflated to 140 after accounting for uncertainty in *RR* and f_T, with mean points per targeted sample of 4.76 with 95% confidence interval from 2.53 to 7.36. These sample sizes were intended to detect one or more positive samples if the prevalence was 0.5% or greater. In the sampled population, 5.08% of the target population was Johne's disease-positive compared to 1.85% of the nontarget population, for an overall prevalence of 1.87%.

The required number of targeted samples represented 23% of the number of simple random samples (Table 3). Despite fewer samples, targeted sampling detected one or more positive samples at nearly the same frequency as simple random sampling. However, the estimate of prevalence from targeted sampling was biased downward from 1.87 to 1.15%. Furthermore, the distributions of estimated prevalence differed (Fig. 2) and the bias ratio

Table 3
Bootstrap sampling results from second populations for specific animal diseases.

Statistic	Simple random sampling	Targeted sampling
Scrapie		
Sample size	1496	405
Samples as proportion of simple random sample	100%	27%
Probability of detecting one or more positive samples	98.3%	98.8%
Estimated prevalence, $\hat{P}$	0.27%	0.29%
Standard error of $\hat{P}$	0.13%	0.15%
Bias ratio	–	0.15
Johne's disease		
Sample size	598	140
Samples as proportion of simple random sample	100%	23%
Probability of detecting one or more positive samples	100.0%	99.9%
Estimated prevalence, $\hat{P}$	1.87%	1.15%
Standard error of $\hat{P}$	0.55%	0.56%
Bias ratio	–	1.27
Neospora caninum		
Sample size	298	122
Samples as proportion of simple random sample	100%	41%
Probability of detecting one or more positive samples	82.3%	95.3%
Estimated prevalence, $\hat{P}$	0.58%	0.94%
Standard error of $\hat{P}$	0.44%	0.76%
Bias ratio	–	0.47

Table 2
Inputs and calculations necessary to derive simple random and targeted sample sizes based on input data.

Model value	Symbol	Formula	Scrapie	Johne's disease	Neospora
Fraction of population in target group	f_T		Beta(1445, 4818)	Beta(627, 33029)	Beta(1199, 33042)
Risk ratio	*RR*	P_T/P_o	Pert(1, 20, 139)	Pert(1.97, 4.66, 10.40)	Pert(0.65, 2.8, 9.41)
Design prevalence	P	$f_T P_T + (1 - f_T)P_o$	0.20%	0.50%	1.0%
Prevalence in targeted group*	P_T	$RR \times P_o$	0.74%	2.18%	2.63%
Prevalence in non-targeted group*	P_o	$P/[f_T(RR - 1) + 1]$	0.04%	0.47%	0.94%
Sampling error	α		0.05	0.05	0.05
Number of simple random samples	n	$\ln(\alpha)/\ln(1 - P)$	1496	598	298
Points per targeted sample*	γ	P_T/P	3.72	4.36	2.63
Number targeted samples (no uncertainty)		n/γ	377	125	113
Number of targeted samples (uncertainty)	n_T	n/γ	405	140	122

* Estimated value using best estimates for other parameters.

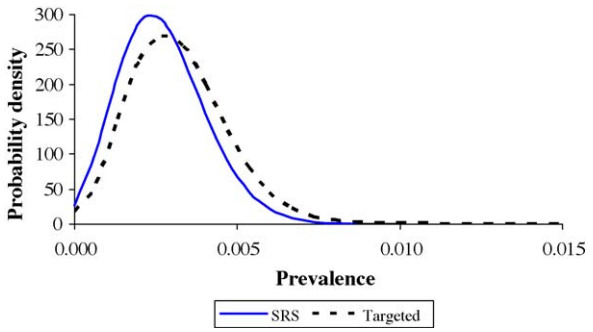

Fig. 1. Distribution of estimated prevalence of scrapie in US sheep population based on targeted sampling compared to simple random sampling.

was 1.27, which greatly exceeds the generally accepted limit of 0.5 and indicates that confidence intervals are unlikely to contain the true prevalence value.

3.3. *N. caninum*

The sample size for simple random sampling in this situation was 298 (Table 2). The targeted sample size based on most-likely values for *RR* and f_T (i.e., 2.8 and 3.5%) was 113. The derived targeted sample size was inflated to 122 after accounting for uncertainty in *RR* and f_T using Monte Carlo methods (Eq. (1.3)), with mean points per targeted sample of 3.19 with 95% confidence interval from 1.10 to 5.72. These sample sizes were intended to detect one or more positive samples if the prevalence was 1.0% or greater. The true characteristics of the sampled population were as follows: 2.47% of the target population were *Neospora*-positive compared to 0.52% of the nontarget population, for an overall prevalence of 0.58%.

The required number of targeted samples represented 41% of the number of simple random samples (Table 3). Despite fewer samples, targeted sampling had a higher probability of detecting one or more positive samples compared to simple random sampling. The targeted sampling estimate of prevalence was biased upward from 0.58 to 0.94% though the distributions of estimated prevalence for each method were relatively similar

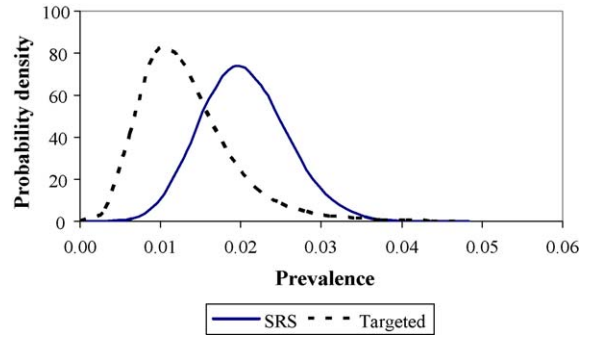

Fig. 2. Distribution of estimated prevalence of antibody to *M. paratuberculosis* (Johne's disease) based on targeted sampling compared to simple random sampling.

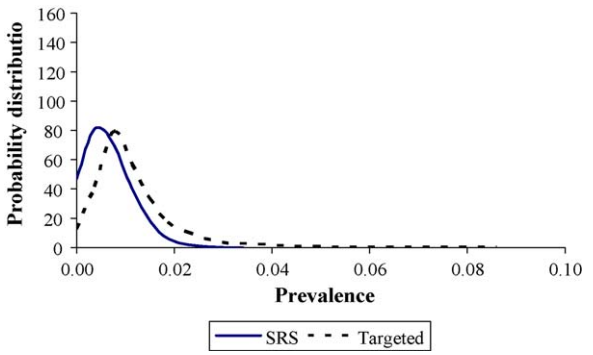

Fig. 3. Distribution of estimated prevalence of antibody to *Neospora caninum* based on targeted sampling compared to simple random sampling.

(Fig. 3). The bias ratio was 0.47, which we consider an acceptable level.

4. Discussion

This report's objective was to evaluate the performance of targeted sampling using specific examples. These examples were based on actual data generated from epidemiologic studies and therefore incorporated epidemiologic uncertainties that are common when planning, implementing, and analyzing surveillance data. In the animal disease scenarios selected, we knew the true population characteristics of the populations sampled but, in reality, we would never know these characteristics. Given these true characteristics, we compared the performance of targeted sampling to simple random sampling.

In the scrapie example, if we sampled a population with 0.20% prevalence, then both simple random and targeted sampling would have equivalent probabilities of detecting one or more positive samples because that was how sample sizes were derived (i.e., both simple random and targeted sample sizes were determined to be 95% confident in detecting one or more positive samples if prevalence is 0.20%). Relative to the true prevalence in the population, simple random sampling was unbiased but targeted sampling is biased slightly upwards. This result is predicted for targeted sampling when based on uncertain *RR* and f_T. Nevertheless, the bias is small relative to inferences about prevalence of scrapie in the population. The similar standard errors generated from each sampling design are also consistent with theory. Despite the uncertainty about *RR* in the reference population of the scrapie study, little bias or loss in statistical precision occurred by using targeted sampling in this example. Concern about these deviations from the performance of simple random sampling should be placed in context with the substantial reduction in samples needed to accomplish targeted sampling in this example. These results support the contention that there is very little difference in the performance of targeted sampling relative to simple random sampling when highly representative and valid estimates of *RR* and f_T are available for use during the study design phase.

In the Johne's disease scenario, the failure of targeted sampling to outperform simple random sampling is the result of overestimates of *RR* and f_T from the study of the reference population relative to the sampled population. Relative to the true prevalence in the sampled population, simple random sampling was unbiased but targeted sampling was biased downwards. In contrast to the scrapie example, where sampling was from essentially the same population, sampling in this Johne's disease example was designed based on results from one population with sampling performance evaluated in a different population at a later time period. In this case, the second population's true *RR* and f_T with diarrhea were less than the central tendencies of *RR* and f_T in the first population that was used to design the sample plan. Consequently, the number of points per targeted sample estimated from the reference population was, on average, larger than the true points per sample in the sampled population. Larger numbers of points per sample implied a larger denominator in the estimator for targeted sampling, so the estimated prevalence in the sampled population (1.15%) was less than the true prevalence (1.87%).

The standard error for simple random sampling is unbiased (e.g., $SE(\hat{p}) = \sqrt{(\hat{p}(1-\hat{p}))/n} = \sqrt{(1.86\% \times 98.14\%)/598} = 0.55\%$) and slightly less than that for targeted sampling. The uncertainty in *RR* and f_T slightly inflated the standard error for targeted sampling. The biased estimator of prevalence for targeted sampling illustrated what can happen when the model used to design a sampling plan is not consistent with the sampled population. Nevertheless, the bias in this Johne's disease situation was not particularly concerning; although the proportional magnitude of bias is −38% ((1.16% − 1.86%)/1.86%), this only underestimated true prevalence by 0.70%. This level of bias might be more significant in other situations.

The confidence level of detection and prevalence estimates in this Johne's disease example illustrated an important consideration for targeted sampling. Unless there is very strong justification for assuming the sampled population is similar to the population(s) used to estimate *RR* and f_T, then the uncertainty for these parameters should be modeled using uninformed distributions. In this example, if we had assumed wider distributions for *RR* and f_T, then a larger targeted sample would have been derived. Furthermore, if the uncertainty distributions for *RR* and f_T were less informed, it is possible that the designed targeted sample could perform similarly to the simple random sample. For example, if we assumed *RR* and f_T were distributed $RR \sim Uniform(1, 10)$ and $f_T \sim Uni\text{-}Uniform(0.005, 0.05)$, then the targeted sample size would have been 189 (still just 32% of the simple random sample size) and the confidence in detection is similar for targeted sampling compared to simple random sampling (e.g., 100% vs. 99.99%). In this case, the mean points per targeted sample were 4.79 with expanded confidence interval from 1.22 to 8.40. Furthermore, the targeted sampling estimate for prevalence is 1.40% with a larger standard error (e.g., 1.04%); this estimate remains biased (bias ratio = 0.45) but the magnitude of the bias was reduced by including more uncertainty about *RR* and f_T during the initial planning stage.

In the *N. caninum* scenario, *RR* was obtained from a nonrepresentative reference population, with different diagnostic tests used including different cutpoints for determining positivity. Despite these differences, targeted surveillance performed relatively well. With 41% of the simple random sample size, the targeted sample was more likely to detect one or more positive samples (95% compared to 82%), though it overestimated the population prevalence (0.94% compared to 0.58%). In this case, the *RR* from the reference population underestimated *RR* in the sampled population, leading to lower points per targeted sample which deflated the denominator, resulting in an inflated prevalence. This indicates the potential for distortion of estimators even with relatively low bias ratios.

All the examples demonstrate that increasing uncertainties about *RR* and f_T cause the targeted sample size to increase, but the improvement in estimation from including more uncertainty seems to more than balance the increase in sample size for the Johne's scenario. It should be noted, however, that stretching uncertainty about *RR* to values below 1 could dramatically increase the targeted sample size. Nevertheless, if *RR* can plausibly be less than 1, then the appropriateness of the proposed characteristic should be questioned; a characteristic with a risk ratio of one does not constitute a justification of using targeted sampling.

Ultimately, increasing uncertainties about *RR* and f_T may translate into values of n_T that are nearly equivalent to the simple random sample size. In such cases, the statistical and technical complexities of targeted surveillance suggest it is not preferred to a simple random sample. This is especially true if the marginal cost of conducting targeted sampling is greater than this cost for simple random sampling. Furthermore, the marginal cost of targeted sampling may increase with increasing samples (e.g., as the pool of potential targeted samples becomes more difficult to identify) and the economic advantage of targeted sampling may dissipate in such cases. As with any surveillance planning, the economics of sampling is an important factor in ultimately deciding which strategy to employ.

The scenarios in this report ignore the effects of sensitivity and specificity on the processes of designing and analyzing targeted surveillance. Formal incorporation of these parameters and their uncertainty into these processes is a next step in the development of targeted surveillance. Similarly, these methods can be extended to multiple targeted subpopulations, but the mathematics are more complex.

5. Conclusion

In summary, targeted surveillance may be appropriate for many animal health surveillance activities. This approach to surveillance is highly appropriate when one has representative and justifiable estimates of key epidemiologic parameters. If one is less certain of the

epidemiologic parameters *RR* and f_T, however, a rational decision is to reflect this uncertainty across the full range of plausible values. Although this uncertainty necessarily inflates the targeted sample size, it is expected to improve the validity of estimates made from targeted surveillance.

Acknowledgments

The authors would like to acknowledge the USDA-APHIS-Veterinary Services for funding this study.

References

Cannon, R.M., 2002. Demonstrating disease freedom—combining confidence levels. Prev. Vet. Med. 52, 227–249.

Hernandez, J., Risco, C., Donovan, A., 2002. Risk of abortion associated with *Neospora caninum* during different lactations and evidence of congenital transmission in dairy cows. JAVMA 221, 1742–1746.

Minnesota Board of Animal Health, 2006. Minnesota Bovine Tuberculosis Management Plan. Minnesota Board of Animal Health, Minnesota Department of Agriculture, Minnesota Department of Natural Resources, and U.S. Department of Agriculture, August 11, 2006. http://www.bah.state.mn.us/diseases/tuberculosis/Minnesota%20Bovine%20Tuberculosis%20Management%20Plan.pdf, accessed 8-29-07.

Prattley, D.J., Cannon, R.M., Wilesmith, J.W., Morris, R.S., Stevenson, M.A., 2007. A model (BSurvE) for estimating the prevalence of bovine spongiform encephalopathy in a national herd. Prev. Vet. Med. 80, 330–343.

Särndal, C.E., 1978. Design-based and model-based inference in survey sampling. Scand. J. Stat. 5, 27–52.

Särndal, C.E., Swensson, B., Wretman, J., 1992. Model Assisted Survey Sampling. Springer-Verlag, New York.

Stärk, K.D., Regula, G., Hernandez, J., Knopf, L., Fuchs, K., Morris, R.S., Davies, P., 2006. Concepts of risk-based surveillance in the field of veterinary medicine and veterinary public health: review of current approaches. BMC Health Service Res. 6, 20.

USDA-APHIS-VS, 1996. Part I: Reference of 1996 Dairy Management Practices. USDA-APHIS-VS, CEAH, National Animal Health Monitoring System, Fort Collins, CO, #N200.696.

USDA-APHIS-VS, 1997. Johne's Disease on U.S. Dairy Operations. USDA:APHIS:VS, CEAH, National Animal Health Monitoring System, Fort Collins, CO, #N245.1097.

USDA-APHIS-VS, 2003. Distribution of genotypes at codon 171 in US sheep. USDA-APHIS-VS, CEAH, National Animal Health Monitoring System, Fort Collins, CO, #N413.1203.

USDA-APHIS-VS, 2004. SOSS. Phase II: Scrapie: Ovine Slaughter Surveillance Study 2002–2003, #N419.0104, 30 pp.

USDA-APHIS-VS, 2005. Johne's Disease on U.S. Dairy Operations, 2002. USDA:APHIS:VS: CEAH, National Animal Health Monitoring System, Fort Collins, CO, #N427.0205.

USDA-APHIS-VS, 2006. Bovine Spongiform Encephalopathy (BSE) Ongoing Surveillance Plan, July 20, 2006. http://www.aphis.usda.gov/newsroom/hot_issues/bse/downloads/BSE_ongoing_surv_plan_final_71406%20.pdf, accessed 8-29-07.

Williams, M.S., Ebel, E.D., Wells, S.J., 2009. Population inferences from targeted sampling with uncertain epidemiologic information. Prev. V. Med., Feb 4 [Epub ahead of print].